

Abstract

A software method and computer system using said method are disclosed for efficiently reallocating unused system resources in a system in which unused system resources are initially allocated among groups that may become active or inactive. The method reallocates unused resources among active groups by scaling up these groups' initial entitlement shares, subject to maximum caps that may apply. For each group, the method calculates a scaling ratio equal to the cap divided by the entitlement, and then sorts the groups according to their scaling ratios. The method processes each group in increasing order of their scaling ratios and reaches a maximal share of system resources to be allotted to each group, where the maximal share for an active group is between, or equal to, the group's initial entitlement value and its cap. Inactive groups have maximal shares equal to zero, reflecting that they receive no system resources. As the method processes the groups, it attempts to simply scale each active group by its scaling ratio. If scaling all unprocessed groups by the current group's scaling ratio would exceed the unallocated resources, then the remaining groups are scaled by the remaining resources divided by the sum of the entitlements of the remaining groups. Because the groups are sorted in order of increasing scaling ratios, the remaining active groups may be scaled as described without exceeding their caps.